DOCUMENT RESUME

ED 270 842 EA 018 516

AUTHOR Murnane, Richard J.; Cohen, David K.
TITLE Merit Pay and the Evaluation Problem: Understanding Why Most Merit Pay Plans Fail and a Few Survive.
SPONS AGENCY Stanford Univ., Calif. Inst. for Research on Educational Finance and Governance.
PUB DATE Aug 85
NOTE 39p.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Accountability; Educational Economics; Elementary Secondary Education; *Incentives; *Merit Pay; *Public School Teachers; School Community Relationship; School Districts; School Support; Teacher Administrator Relationship; *Teacher Evaluation; *Teacher Motivation

ABSTRACT

Merit pay for teachers is often proposed as a compensation system that will enable public schools to motivate teachers to pursue organizational goals. Yet the promise of merit pay is dimmed by knowledge of its history: most attempts to implement merit pay for public school teachers over the last 75 years have failed. The first part of this paper uses microeconomics to explain the failures of most merit pay plans, and to show that merit pay, even taken on its own terms, does not provide a solution to the problem of how to motivate teachers. The second part investigates why merit pay survives in a very few school districts. The analysis is based on interviews with teachers and adminstrators in six school districts with enduring merit pay plans. The interviews indicate that in these exceptional districts, merit pay contributes to solving problems quite different from the problem of motivating teachers. These problems include (1) how to support good teachers who differ in their relative needs for income and free time; (2) how to encourage meaningful dialogue between teachers and administrators on issues such as instructional evaluation, and (3) how to build community support for schools. References are included. (TE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED270842

MERIT PAY AND THE EVALUATION PROBLEM:
UNDERSTANDING WHY MOST MERIT PAY PLANS FAIL AND A FEW SURVIVE

Richard J. Murnane* and David K. Cohen**

* Graduate School of Education, Harvard University

** Michigan State Univerasity

August 1985

The research on which this paper is based was supported by the Institute for Finance and Governance, School of Education, Stanford Univerasity. We would like to thank Katherine Jamentz, Patrick Murnane, and Niall Nelson for able research assistance, and Chris Argyris, Estelle James, Susan Johnson, Charles Lindblom, Giandomenico Majone, Richard Nelson, Robert Pollak, David Stern, Dvora Yanow and participants in workshops at Boston Univerasity, Harvard Univerasity, and the Univerasity of Pennsylvania for helpful comments on earlier drafts. We would especially like to thank Edward Pauly, who read several drafts of the paper, and provided many important ideas.

EA 018 516

I. INTRODUCTION

Designing a compensation system that provides strong incentives for employees to pursue organizational goals vigorously is a challenge every organization faces. Merit pay for teachers is often suggested as a compensation system that will enable public school systems to meet this challenge. Yet the promise of merit pay is dimmed by knowledge of its history: most attempts to implement merit pay for public school teachers over the last 75 years have failed.

The first part of this paper uses microeconomics, the intellectual home of merit pay, to explain the failures of most merit pay plans. We show that merit pay, even taken on its own terms, does not provide a solution to the problem of how to motivate teachers.

The second part of the paper investigates why merit pay survives in a very few school districts. Our analysis is based on interviews we conducted with a great many teachers and administrators in six school districts with enduring merit pay plans. We explain that in these quite special districts, merit pay contributes to solutions to problems quite different from the problem of motivating teachers.

II. COMPENSATION OF PUBLIC SCHOOL TEACHERS

A. Uniform Salary Scales

More than 99 percent of the public school teachers in the United States work in districts that employ uniform salary scales. Under such contracts, a teacher's salary is determined exclusively by educational credentials and years of teaching experience. All teachers with the same credentials and experience receive the same salary, irrespective of subject specialty or perceived performance. Typically each school district sets its own salary scale, or negotiates it with the local teachers' union through collective bargaining.

The limitations of uniform salary scales have been well documented: no financial reward for superior performance; no financial penalty, short of dismissal, for inferior performance (Hanushek, 1981). Many critics of uniform salary schedules argue that improvement in the quality of education offered by public schools requires a change from uniform salary schedules to a compensation scheme that bases individual teachers' compensation on their performance in teaching their students, as measured either by student test score gains, or by supervisors' evaluations of teacher actions in the classroom. Such performance-based compensation plans are typically called merit pay.

B. A Brief History of Merit Pay

Merit pay in an old idea. In 1918, 48 percent of United States school districts sampled in one study used merit pay (Evendon, 1918, as reported in Johnson, 1984). Little is known about these early plans, except that most did not last. In 1923, the National Education Association (NEA) reported that 33 percent of sampled districts used merit pay. In 1928, a subsequent NEA survey reported that 18 percent of districts surveyed awarded merit pay.

Interest in merit pay waned during the 1940s and early 1950s, as the vast majority of public school districts adopted uniform salary schedules. Between 1939 and 1953, the percentage of school systems in cities with more than 30,000 population that used merit pay fell from twenty to four (Porwoll, 1979, p.26).

Sputnik rekindled interest in merit pay. During the 1960s, approximately 10 percent of United States school districts had merit pay plans. Most of these plans fared no better than their predecessors. By 1972, the percentage of districts using merit pay had fallen to 5.5 percent (Porwoll, 1979). A 1978 survey of the 11,500 U.S. school districts with enrollments of 300 or more students found only 115 with merit pay plans (4 percent of the districts that responded to the survey; 1 percent of districts to whom the questionnaire was sent). Moreover, the majority of districts that reported having tried and dropped merit pay indicated that their plans lasted less than five years (Porwoll, 1979, p.41).

Thus, the history of merit pay suggests that interest in paying teachers according to merit endures, but attempts to use merit pay do not. Moreover, the geographical and temporal nature of the evidence indicates that teacher union resistance cannot account for the demise of most merit pay plans. We must search for other explanations. We believe that the most powerful ideas for understanding merit pay can be found in the literature of economics, which has generally been thought to provide the intellectual basis for merit pay. Specifically, we turn to economists' analyses of employment contracts, a growing field within microeconomics.

III. WHY MOST MERIT PAY PLANS FAIL

A. The Contracts Literature: A Framework for Analysis

One branch of microeconomics, which we will call the contracts literature, examines the cost and benefits associated with using different types of employment contracts to compensate workers engaged in particular kinds of production activities. The following assumptions underlie this literature:

1. Workers' preferences are not completely consonant with the employing organization's goals. If there are no adverse consequences for them, workers prefer to work less hard than the organization would like.

2. Monitoring the output of individual workers or the actions of individual workers is costly.
3. Imperfect monitoring will induce workers to attempt behavior that makes them appear productive relative to other workers, but in fact is contrary to the goals of the organization. Williamson (1975), an important contributor to the contracts literature, labels this behavior opportunistic, and defines it as self-interest seeking with guile.

As seen from the perspective of the contracts literature, the type of employment contract an organization should adopt depends on the type of work employees of the organization do. The reason is that the cost of evaluating workers' output, the cost of evaluating workers' actions, and the potential for opportunistic behavior all depend on the nature of the production activity.

The framework provided by the contracts literature is helpful in analyzing merit pay for three reasons. First, this literature takes seriously the evaluation problem. By this, we mean that it explicitly acknowledges that evaluating workers' performances is costly for management, and that imperfect evaluations, defined as less than perfect knowledge of all worker actions, may elicit unpredicted and potentially destructive responses from workers. It is this evaluation problem that has plagued most attempts to introduce merit pay to public education.

Second, the contracts literature focuses attention on the nature of the production activity in which workers are engaged.

It explains why an analysis of the production activity provides the best clues to the responses that particular compensation plans will elicit. We will argue that compelling explanations for the failure of most merit pay plans must focus on the nature of teachers' work.

Third, the contracts literature emphasizes the importance of tradeoffs between the gains from providing incentives for employees to work hard and the costs of various ways of evaluating workers' contributions. Implicit in this emphasis on tradeoffs is the often neglected recognition that a merit pay system that brings about modest increases in teachers' effort levels might not be worthwhile if the costs of the measures taken to evaluate workers' contributions are extremely high.

In the next sections we use the framework provided by the contacts literature to explain the evidence on the two main kinds of performance-based compensation plans, known as "new style merit pay" and "old style merit pay."

B. "New Style Merit Pay": A Piece-Rate Compensation System

One type of merit pay, which has been called "New Style Merit Pay,"[1] and payment by results,[2] bases individual teachers' merit pay bonuses on the test score gains of their students. The attractiveness of this strategy is that the evaluation problem is solved by actually measuring certain dimensions of each teacher's output, and thereby avoiding the subjectivity of the evaluations under "old style merit pay," where bonuses are based on

supervisors' evaluations of teachers' actions. There are only a few documented cases of school districts that have used "new style merit pay", although merit pay plans that compensate teachers on the basis of student test score gains have recently been supported by several state legislatures.[3] In this section we show that "new style" merit pay is very much like what economists know as a piece-rate compensation system, and that teaching does not satisfy the conditions under which this type of compensation system is efficient.

Approximately 30 percent of U.S. workers in manufacturing are employed under piece-rate contracts, the most common form of payment by results (Pencavel, 1977; Seiler, 1984). Piece-rate contracts work well when the true contribution of the individual worker to the firm's output can be measured at relatively low cost. Commercial laundries' contracts with workers who iron shirts provide an example. The number of shirts ironed provides a good measure of the worker's contribution to the firm. Consumer complaints provide a check on quality. Multiple dimensions of output can be managed by providing a schedule of piece rates for different types of clothing.

Piece-rate contracts do sometimes elicit opportunistic behavior. For example, workers may neglect the maintenance of the machines on which they work, since they are not rewarded for machine maintenance.[4] For many types of work, however, the costs of such opportunism are outweighed by the advantages that piece-rate contracts have over contracts that attempt to control opportunism by monitoring worker actions. In particular,

piece-rate contracts provide strong incentive for workers to find the most rapid way to iron shirts. High productivity results in immediate rewards; shirking results in immediate penalty.

Why haven't merit pay plans that compensate teachers on the basis of their output, measured, for example, by student test score gains, become popular? One reason concerns the nature of the incentives that such a compensation system provides. Any explicit list of pay rates for specific levels of student test score gains (economists would refer to such a list as a payment algorithm) creates a specific price--a piece-rate--for each student's test score gain in each subject area. For example, an algorithm that bases compensation solely on average reading score gains implicitly places a zero price on student gains in other subject areas. Moreover, it places an equal weight on each student's gain. If teacher time is viewed as a private good (time spent with one student reduces time available for other students), then this algorithm creates incentives for teachers to allocate time so that the last unit of time spent with any child yields the same expected test score gain. This means that there are incentives for teachers to minimize the time they spend with children whose test scores will not respond to modest increases in attention.

There is limited evidence that teachers do respond to payment by results by allocating their time to specific subject areas and children. For example, in the middle of the 19th century in England, elementary school teachers worked under a payment by results plan that based their compensation on the number of

children who acquired a set of narrowly defined skills. This led to a narrowing of the curriculum to exclude all nontested skills, including many skills that were perceived to be important, but were difficult to test (Coltham, 1972).

Other evidence comes from the performance contracting experiments sponsored by the Office of Economic Opportunity in the early 1970s. In these experiments, private firms provided reading instruction to public school children, with the firm's compensation dependent on student test score gains. In at least one of the sites, teachers concentrated their time on children in the middle of the test score distribution, neglecting children at the top of the distribution, who would advance well on their own (test score gains above a threshold were not rewarded), and children at the bottom of the distribution, whose test scores would not respond to modest additional amounts of teacher time (Gramlich and Koshel, 1975).

Why are the responses of teachers to the incentives implicit in the test score algorithm problematic? If different time allocations among subjects and students are desired, why not simply alter the algorithm, for example, by giving weight to skill development in more skill areas, and perhaps by weighting achievement gains of some children more than those of others? There are two problems with this solution: lack of consensus about the appropriate weights, and the nature of teachers' work. Consider each problem in turn.

Most policy debates about public education avoid the divisive topic of weights--at its core, a discussion about whose education

matters the most. Instead of explicitly debating what the weights should be, it is common in public education to delegate resource allocation decisions to teachers and administrators, with the inoperable admonition that they provide every student with the opportunity to fulfill his or her potential. Such delegation is not consistent with the design of contracts that pay teachers on the basis of their output.

If the public schools' lack of consensus on weights were the only problem in paying teachers on the basis of their students' progress, one would expect to see more extensive use of such compensation schemes in private schools, where, presumably, family choice leads to greater agreement on school goals. The limited available evidence suggests that performance based pay for teachers is relatively rare in private schools, however. In 1983, only seven percent of Catholic high schools used any form of merit pay, and none of those schools base pay differentials on student test score gains (The Catholic High School: A National Portrait, 1985).

Why aren't teachers paid on the basis of their students' test score gains, even in organizations in which there is relatively high consensus on goals, union power is negligible, and contract form is management's prerogative? We believe that the answer concerns the nature of the work in schools. Even in schools in which there is a high level of consensus on goals, the goals are multidimensional--for example, raise the average reading level in each class, and eliminate drugs and violence from the school. While it may be reasonable to attribute the average reading gain

of a group of students to a particular teacher working with students behind a closed classroom door, it is not possible to measure each teacher's contribution to the second goal. In fact, eliminating violence and drugs from a school requires that teachers open their classroom doors and work as a team to monitor students' actions while not in class. If teachers' pay is based solely on success in raising reading scores, teachers have strong incentives to keep their classroom doors closed, and neglect the teamwork outside of classrooms that contributes to accomplishment of the second goal. Moreover, the strategy used in some laundries to combat this form of opportunism, namely, hiring workers to tend to the tasks neglected by piece-rate workers (for example, maintaining the machinery), does not work well in schools. As many schools have found, teachers, who work with students every day in class, and who know students' names and their personalities, are typically more effective in eliminating drugs and violence from a school than specialized security officers are.

School principals, as well as teachers, realize that much of the important work in schools must be done by teachers working together and that compensation algorithms that reward only those dimensions of performance for which each teacher's contribution can be measured create perverse incentives. This may explain why paying teachers on the basis of their students' test scores is extraordinarily rare in American education.

It is important to note that our explanation rests on the nature of teachers' work and the incentives that piece-rate

compensation schemes provide. This is quite different from the typical objection to "new style merit pay", which emphasizes the inadequacies of standardized tests. While it is true that standardized tests of, say, students' reading skills often do not provide measures of students' true skills, and consequently of the fruits of teachers' work, the limitations of tests are not the fundamental problem with "new style merit pay." Even if tests were developed that measured quite accurately students' skills in particular subject areas, incentives to allocate time strategically to particular students and particular subject areas, and to neglect aspects of the job such as eliminating drugs and violence, would remain.

C. "Old Style Merit Pay"

The significance of teamwork and the presence of school principals who have direct supervisory functions suggest the possibility of basing teachers' compensation on principals' evaluations. In fact, such "old style merit pay" is the common model for merit pay. In this section, we explore the extent to which the framework provided by the contracts literature helps us to understand why most experiments with "old style merit pay" have failed.

The lessons from the contracts literature regarding the conditions under which it is efficient to base the compensation of individual workers on supervisors' assessments of their actions can be summarized as follows. Merit pay is efficient

when the nature of the activity in which workers are engaged is such that supervisors can provide relatively convincing answers to the following two questions posed by workers:

1. Why does worker x get merit pay and I don't?
2. What can I do to get merit pay?

Unloading boxes from a truck is often suggested as an activity where supervisors can answer workers' questions about performance-based pay differentials. Supervisors can state that worker x was paid more than other workers because he carried two boxes at a time, while other workers carried one at a time. Workers are likely to accept this answer because they recognize that carrying two boxes at a time is in fact productive. They also recognize that the nature of the activity gives worker x few possibilities for opportunistic behavior--that is, for actions that make him appear productive, but in fact do not contribute to the unloading. Supervisors can answer workers' second question by stating that they can earn higher pay by carrying two boxes at a time. Workers are likely to find this answer acceptable because the required action is something they can do, if they choose to.

The nature of teachers' work is very different from jobs like unloading a truck. As is the case with all jobs, some teachers are more effective than others--hence, the call for merit pay. However, most analysts agree that there are no clearly defined behaviors that consistently result in high performance.[5] In other words, there is no analog to carrying two boxes on every trip.

One consequence of the imprecise nature of the teaching activity is that supervisors cannot answer convincingly teachers' questions about why teacher x got merit pay and they did not. As one of the administrators we interviewed commented:

> ....I know who the good teachers are. They're so and so, so and so, and so and so. Why are they good teachers? Well, I don't know; they are just good teachers; but I know who they are...

Many teachers who are denied merit pay find this answer unsatisfactory. One reason is that they are aware that the nature of teaching, with its closed classroom doors, and the networks of relationships among teachers and between teachers and parents, provides great potential for opportunistic behavior. In other words, there are many things that a teacher could do to impress a principal and to suggest that he was more effective than his colleagues. Examples include using friendships with parents to spread rumors about other teachers' incompetence, and not sharing materials that could help other teachers. Thus, teachers have reason to question whether merit pay went to teachers who were in fact the most productive or to those who were most facile in impressing supervisors.

A second consequence of the imprecise nature of teaching is that supervisors cannot answer convincingly teacher's second question: what can I do to earn merit pay? In other words, they cannot suggest specific actions that the teacher can undertake that both the teacher and the supervisor recognize will enhance the teacher's effectiveness. Without an unequivocal answer to

this second question, teachers may have little incentive to change their behavior in pursuit of higher income. Worse yet, teachers may learn the lesson that concealing their problems and playing up to evaluators is what the organization rewards--dramatically complicating managers' evaluation problem.

In effect, the lesson from the contracts literature is that the problems with "old style merit pay" are more fundamental than careless implementation, or inadequate training of evaluators--to name but two of the explanations often given for the failure of merit pay plans. The problem lies in the nature of the teaching activity itself. Specifically, it is the lack of a blueprint for effective teaching that prevents supervisors from providing convincing answers to teachers' questions about why teacher x got merit pay and they did not, and what they can do to earn merit pay.

What the contracts literature does not reveal is exactly what problems arise under merit pay that has led most school districts to drop this type of compensation system after a brief trial. While our field work was not designed to address this question, we did learn some interesting things about the problems merit pay causes, particularly from teachers and administrators in two districts that have had merit pay for more than twenty years, and have altered their plans several times to deal with perceived problems. The comments of these participants are informative in understanding what happens when supervisors cannot answer teachers' questions about why some teachers get merit pay and others don't.

One theme that ran through our interviews is administrators' perception that many teachers who received evaluation ratings lower than they felt was just responded by working less hard. As one teacher told an administrator: "If that's all you care, then that's all you'll get."

What constitutes a just rating from a teacher's point of view? We heard frequently that the absolute floor was the rating the teacher had received in the last rating period. Rarely do discussions of merit pay focus on the significance of the repetitive nature of the evaluation process. Yet teachers see their "merit" ratings in terms of what they and others have been told by supervisors in the past.

The rising expectations and the pressures from teachers for higher ratings over time were recurrent themes in the comments of admininistrators. Moreover, several administrators cited negative consequences that arose from giving a teacher a lower rating than the teacher had received in the past. One principal mentioned that when he gave an older teacher with no better than adequate performance a rating lower than his predecessor had given her, the teacher, who had planned to retire at the end of the year, became so incensed that she postponed her retirement for two years. The key point here is that even an evaluation system that produces valid and reliable performance ratings is not enough to guarantee the success of merit pay. If teachers feel that the ratings are unjust and evaluators cannot convince

them to the contrary, their reactions to the ratings may undermine the education students receive.

Several school principals commented that prior to the introduction of merit pay, they often gave teachers ratings higher than the teachers actually deserved, and encouraged the teachers to live up to the high ratings. The principals reported that this was the most effective strategy for stimulating many teachers to improve their performance. While this evaluation strategy produced ratings that were not objectively valid, the principals felt it promoted their primary goal of stimulating teachers to work hard in helping children to learn. In particular, this approach allowed administrators to focus on the specific problems a teacher was struggling with, without an atmosphere of concealment.

Many of these same principals decried the use of merit pay because it reduced their ability to pursue the strategy of encouraging teachers through high ratings. Under merit pay, principals were under pressure from superintendents, who must defend the cost of merit pay to their school boards, to keep the number of top ratings low and to provide a strong defense of the top ratings they did give. At the same time, principals felt under pressure from teachers to provide explanations for why they had not been given the top rating while the teacher in the next classroom had. As one principal stated: "merit pay turns my job from being a coach into being a referee." And he implied that his teachers no longer saw him as a helpful coach, but as a

critical referee--and this threatened his ability to motivate his teachers to high effort levels.

Some readers may conclude that one of the benefits of merit pay is that it pressures principals into actually evaluating teachers objectively, one of the most important parts of their job. There is something to this argument. Most principals whom we interviewed certainly felt this pressure. There is more to be said, however. School principals need teachers. Principals' primary job is to ensure that the children who pass through their school learn as much as possible. Yet they don't teach any students. Teachers do the teaching, and the principals' success depends to a large extent on their ability to encourage teachers to work hard and work together. When a principal gives a teacher a lower evaluation than the teacher had previously received, the teacher may lose some money. But the principal may lose the cooperation that he needs to make the school work. Our field notes contain many stories principals told about the distressing consequences of giving teachers' ratings lower than the teachers expected.

In effect, merit pay dramatically increases the tensions surrounding the formal evaluation process, and increases the intensity with which teachers ask why they did not receive the top rating and what they can do to receive a better rating next time. In most experiments with merit pay for teachers, school principals have not been able to respond to these questions convincingly. We believe that the lack of convincing, specific responses to teachers' questions contributes in a central way to

the low morale and "problems of administration" that are cited in a recent survey as the primary reason school districts drop merit pay.[6]

Our evidence leads us to emphasize the importance of the imprecise nature of teachers' work in contributing to the demise of "old style merit pay." One of the readers of an early draft of this paper commented that, if we were correct, and the problem was not simply poor public sector management, we should expect that "old style merit pay" would not be common in for-profit educational institutions. While an in-depth exploration of this proposition was beyond our resources, we did attempt to respond to this comment by learning about the compensation policies of Stanley Kaplan, the largest of the for-profit firms specializing in preparing students to take standardized tests, such as the Scholastic Aptitude Test.

Stanley Kaplan does monitor the performance of its teachers closely, in part by observing them in the classroom, and to an even greater extent, by soliciting student evaluations of each teacher's performance. In fact, Kaplan's students are quick to complain when the quality of instruction does not justify the cost of the course. Kaplan uses the feedback from students in deciding which teachers to dismiss. The firm does not use this information in determining individual teacher's compensation, however. In fact, teachers who work for Stanley Kaplan are paid in much the same way public school teachers are paid. All teachers are paid according to a salary scale that bases compensation on experience--on the number of courses taught.

There are no bonuses for superior performance.

We asked the personnel director of Stanley Kaplan why the firm does not use performance-based pay. Her answer included these points. All Stanley Kaplan teachers are effective; those who are not are dismissed. There are some teachers who are superstars, and the firm has considered paying bonuses to these teachers. The firm has rejected this plan because of its managers' perception that the positive impact of bonuses on the performances of the superstars would be more than offset by negative effects on the performances of effective teachers who don't receive bonuses, don't know why they were passed over, and can't be told how to become superstars.

In the context of our paper, the Kaplan evidence can be interpreted as implying that even when management can make relatively accurate fine-grained distinctions among teachers, management would not be able to convince the merely good teachers of the superior performance of some of their co-workers. As a result, the responses to the pay differentials would not further the goals of the organization. Thus, the imprecise nature of the teaching activity prohibits evaluators from answering the hard questions teachers pose about "old style merit pay" and leads a successful profit-making firm to base compensation on experience. It is important to add that Stanley Kaplan uses evaluation aggressively, even without merit pay, both to dismiss ineffective teachers and to offer useful advice to effective

teachers. This approach is obviously similar to that taken by administrators in many public schools.

IV. WHY DOES MERIT PAY SURVIVE IN SOME SCHOOL DISTRICTS?

A. Our Research Strategy

If merit pay is not an effective strategy for improving teachers' performance, why do merit pay plans survive in a few districts? Are the districts atypical? Are the provisions of the merit pay plans atypical? Did merit pay in these districts help to solve problems other than that of motivating teachers?

We began our search for the answers to these questions by identifying school districts that have used merit pay for a number of years. Two Educational Research Service (ERS) reports were helpful in this regard. The first (Porwoll, 1979) identified 115 school districts in the U.S. that used merit pay in 1978. The second (Calhoun and Protheroe, 1983) reported the results of a survey that inquired whether each of these 115 districts was still using merit pay in 1983, and if not, why not. We used the 47 districts that reported in the 1983 survey that they were still using merit pay as the population from which we selected districts for study.

We had hoped to identify urban districts within this population. The reason is that, since many urban districts are thought to have particularly serious problems with poor teaching

quality and low teacher morale, analysis of enduring merit pay plans in urban districts might provide important insights about the factors that contribute to the success of performance-based contracts for improving teachers' performance. We found no urban districts with long-lived merit pay plans. In fact, we could not find even one documented case of a large, once-troubled school district that had successfully used merit pay to improve its performance. To the contrary, one of the striking aspects of the list of districts with enduring merit pay is the large percentage of very small districts serving relatively homogeneous student populations. Moreover, these districts tend to use very small amounts of money as merit pay bonuses.

We then looked for districts that had used merit pay for at least five years and either used pay differentials of at least $1000, or served more than 10,000 students. We found seven districts that met these criteria. We spent several days in six of these districts interviewing teachers and administrators with the goal of learning how each merit pay plan worked and what teachers' and administrators' reactions to the plans were.

B. Characteristics of the Districts

The six districts that we visited vary in size, from 2500 students to 60,000 students. Three are located in the southwest; one in the northeast; one in the midatlantic region and one in the northcentral region of the country. Two districts have

collective bargaining; the union role in the other four is insignificant.

All of the six districts are considered to be among the best in their geographical areas--places where teachers like to work, and where high housing prices reflect the desirability of the public schools. In evaluating the role merit pay plays in contributing to these districts' accomplishments, it is important to focus first on attributes other than merit pay that the districts have in common.

All of the districts have salary schedules, to which merit pay is added, that are above average for their geographical areas. The high salaries and good working conditions permit these districts to be selective in choosing among applicants for teaching positions. None of these districts adopted merit pay as a response to the idea that there was not enough money to pay all teachers well, so they would at least pay a few good teachers well. In fact, we heard from several administrators comments such as this one: "No merit pay system would ever work without salaries at a point that teachers can live on."

None of these districts use merit pay as a strategy to give negative signals to teachers perceived to be ineffective. However, using evaluation practices quite unrelated to merit pay, these districts do dismiss teachers judged to be incompetent, and are pressured by parents to do so. These practices have not been resisted by teachers' unions in the two districts with relatively powerful unions. The union leaders in these districts stated

that they made sure that due process was observed, but that it was not in the union's interest to protect incompetent teachers.

One lesson to be learned from examining the characteristics of school districts with long-lived merit pay plans is that attractive working conditions may be a prerequisite for the survival of merit pay.

C. Characteristics of the Enduring Merit Pay Plans

Working conditions do not provide the whole answer to why merit pay survives in a few districts. In fact, merit pay has been dropped by a great many districts that appear similar to the districts we studied that do have long-lived merit pay plans. Thus, to explain the survival of merit pay in our districts, we need to look at the plans themselves.

The six merit pay plans that we learned about differ in many respects. However, in every case, the plan incorporates a strategy for dealing with the two questions that many teachers ask about merit pay:

1. Why does teacher x get merit pay and I don't?
2. What can I do to earn merit pay?

The strategies consist of varying combinations of four themes: extra pay for extra work; make everyone feel special; make the program inconspicuous, and legitimation through participation.

1. Extra pay for extra work.

One common theme in the long-lived merit pay plans is that the definition of performance is altered so as to reduce emphasis on classroom teaching, and increase emphasis on completion of tasks outside the classroom. For example, the numerical rating system used by one district to determine merit pay awards gives school and community service the same weight as classroom performance. Another district requires that a teacher complete six outside activities to be eligible for merit pay. As one teacher commented: "This isn't merit pay; it's how you get the yearbook done."

A complementary practice is to make the teacher responsible for documenting that he or she is worthy of merit pay. As part of the merit pay application process in several districts, teachers had to prepare lengthy documents describing their accomplishments and providing evidence in the form of testimonials from colleagues and parents. One teacher commented:

> When I finished this last time, I had a volume no less than three inches thick of evidences, arguments, and materials.

These practices, which we call extra pay for extra work, provide one set of relatively convincing answers to the two questions teachers raise about merit pay. Administrators can clearly state that teacher x received merit pay because he devoted time to organizing a variety of activities and to documenting his accomplishments, both in the classroom and

outside the classroom. If another teacher wants merit pay, he can do these same things.

2. Make everyone feel special.

A second theme is to quietly award merit pay to almost all teachers. This strategy is most pronounced in one district in which a numerical rating system is used to determine whether teachers receive an award of $0, $500, $1000, $1500, or $2000. Teachers whom we interviewed were unaware of the distribution of actual awards, but typically were pleased that they received a substantial award. In fact, every teacher who participated in the voluntary merit pay program (over 90 percent of eligible teachers in the district) received a positive award; 85 percent of the teachers received either $1500 or $2000.

We suspect that the bunching of the ratings at the top of the scale and the relatively small monetary differential between the top two awards is important in minimizing ill-feeling on the part of teachers in schools headed by hard grading principals. In this district, having the principal be a hard grader means that the effective teacher gets a $1500 annual bonus instead of a $2000 bonus.

In effect, the "make everyone feel special" strategy deals with teachers' potentially destructive questions about merit pay by reducing the number of teachers who ask. We find it interesting that the "make everyone feel special" theme was

particularly evident in the two districts in our sample that have had merit pay for more than twenty years.

3. Make merit pay inconspicuous.

In several districts, the design of the merit pay system is such that the incentives are of little interest to a large percentage of the teachers. For example, in one district, eligibility for merit pay requires ten years of service, completion of six activities outside the classroom, and satisfactory performance evaluations. The reward for fulfilling these requirements is $600 (somewhat more, if coupled with advanced degrees). Only 40 percent of the teachers in this district who do fulfill the length of service requirement choose to participate in the voluntary merit pay plan.

In another district, teachers can apply for one of four different award levels, with each level having different requirements. While the award levels are sizable, $1000 for level I, $4000 for level IV, the requirements are so demanding that only twelve percent of the teachers apply for any level (two-thirds of these teachers receive awards). For example, the level IV requirements include a Master's Degree and 30 hours of graduate credits, superior teaching skills, as demonstrated, for example, by "representing the district at the state or national level as a resource person, chairperson, or committee member," and superior professional contributions, as demonstrated, for example, by serving "in an official capacity in the management of

the professional associations or organizations related to a specific field of study" (quotations from school district description of performance based compensation plan). For the vast majority of the teachers in this school district, the financial awards do not justify the extra work.

In all of our districts, merit pay has a low profile. In part, this stems from the perception that merit pay is something almost any teacher could earn, but that the financial rewards do not justify the extra work. Another element is that teachers are urged not to discuss with colleagues either who receives merit pay or the amount of the awards. In these districts, where most teachers like their jobs, the primary effect of secrecy seems to be to reduce teachers' interest in merit pay and thereby to reduce the number of teachers who ask the hard questions about why some teachers get merit pay, and others don't. One could easily imagine, however, that in districts where teachers worked under more difficult conditions, secrecy might have a very different result.

4. <u>Legitimation through participation.</u>

One final attribute of merit pay in our districts concerns the process by which the programs were designed. In all of the districts, teachers played a significant role in the design of the merit pay plans. Moreover, in each of the two districts that have had merit pay for more than twenty years, the system has been revised several times in response to teacher complaints. We

believe that teachers' participation in the design, and redesign, of the merit pay plans contributes to their longevity. One reason is that the process of participation reveals information about teachers' preferences, information that is critical for supervisors to have in order to predict teachers' responses to incentives; but this information is extremely difficult to collect unless teachers volunteer it. Participation gives teachers a reason to volunteer information, and a mechanism through which to do so. A second reason that participation contributes to the longevity of merit pay plans is that it creates the impression that merit pay is not a system thrust upon teachers, but rather one they helped to create. Seen as such, teachers may still ask why some teachers get merit pay and others don't, but the intensity with which they ask the questions is diminished. Teachers recognize that if many of them find the program objectionable, they can change it.[7]

V. A ROLE FOR MERIT PAY?

As the observant reader will already have concluded, merit pay in the districts we studied is not primarily pay dependent on success in teaching students, as measured either by student test score gains, or by supervisors' evaluations of teacher actions in the classroom. This raises two questions: Does merit pay, as it actually operates, have any influence on the performance of the school districts we studied? What does merit pay do that these districts could not do without merit pay?

Merit pay in the districts we studied does not appear to have strong effects on the way teachers teach. This is not surprising given the attributes of the enduring plans. This conclusion is also compatible with the theme of Section III: the nature of teachers' work is such that basing individual teachers' pay on assessments of their performance is unlikely to motivate teachers to work harder.

If merit pay doesn't motivate teachers to work harder, why do a very few districts retain it? Our interviews with teachers and administrators suggest that merit pay helps these six districts solve problems quite different from the problem of motivating teachers. These problems include:

1. how to support a good teachers who differ in their relative needs for income and free time.

2. how to encourage meaningful dialogue between teachers and administrators about difficult issues such as the quality of the evaluation process.

3. how to build community support for the public schools.

In the districts we visited, merit pay contributes to solutions to these problems in the following ways.

1. Extra pay for extra work provides opportunities for teachers with large financial needs to augment their incomes by spending time on school-related activities.

2. The ongoing discussions of how merit pay works, what its problems are, and what changes are needed, provide forums for meaningful dialogue between teachers and administrators

concerning difficult issues such as the nature of the evaluation process.

3. The merit pay plans contribute to the perception that teachers are accountable. As one teacher commented about why members of the community supported merit pay for teachers:

> The people out there who are paying taxes want to make sure that in the area of teacher pay, those who are doing the real work are the ones who get the rewards, above and beyond the standard.

And there is another side to the coin. As one administrator remarked:

> [Merit pay] has meant a lot of money for a lot of teachers that would otherwise not have been provided, knowing the Board of Education.

We do not mean to imply that merit pay is necessary to solve the problems of satisfying teachers' varied needs, encouraging dialogue, and promoting community support for the schools. In fact, a far greater number of school districts solve these problems without merit pay than with merit pay. For example, many districts meet some teachers' needs for extra income through extra pay for extra work without calling this merit pay, and through small grants competitions.[8] Many districts use the collective bargaining process to promote meaningful dialogue between teachers and administrators. Many districts promote community support through volunteer programs, public-private partnerships, and outreach activities.

What we do want to suggest is a different way of looking at merit pay. This is useful because, if the past is any guide to the future, the current, perhaps waning, wave of interest in merit pay will not be the last time that educators feel pressure to adopt this type of compensation plan. In thinking about merit pay in the future, it is useful to ask whether it can play a modest role in contributing to solving the problems of satisfying teachers' varied needs, encouraging meaningful dialogue between teachers and administrators, and promoting community support for the schools. In most school districts, the answer to this question will be an emphatic no. But in a few districts the answer could be a tentative yes. We hope that our work provides some clues about the types of districts where an answer of maybe makes sense, and what types of merit pay plans hold some promise.

There is another theme in our evidence, one that would seem to apply to districts both with and without merit pay. It is the theme of improving teachers' performance through evaluation. In order for evaluation to contribute to the goal of helping teachers improve, it must be carried out by skilled and knowledgeable supervisors in an atmosphere that rewards honesty and cooperation. When teachers who conceal their failings get higher pay than those who do not, the atmosphere for useful evaluation and advice is poisoned. If supervisors are to engage in a useful dialogue with teachers, they must act in a way that is consistent with the sustained nature of their relationship with teachers. Evaluation is a repetitive sequence that creates expectations, memories and sensitivities that can either

contribute to improved performance or, if treated insensitively, undermine it. It was the goal of merit pay's advocates to put the power of money into the evaluation process, as a way to improve teachers' performance. That goal is misguided. But the broader lesson, that school administrators must work to create relationships with teachers in which evaluations contribute to improvement, change, and cooperative problem-solving, is one that must not be forgotten even after the pressures for merit pay dissipate.

FOOTNOTES

1. See Bacharach, Lipsky and Shedd, 1984.

2. See Coltham, 1972.

3. See The Nation Responds: Recent Efforts to Improve Education, 1984, p. 45, for a reference to legislation that provides state financial support to school districts that adopt "new style merit pay" plans.

4. See Pencavel, 1977.

5. See Wise et al., (1984, p. 10) for a discussion of the claims and refutations concerning the role of specific teacher actions in fostering student learning.

6. See Calhoun and Protheroe, 1983.

7. The importance of voice as a mechanism for improving the performance of organizations is elegantly developed by Hirschman, 1972.

8. See, for example, Small Grants for Teachers (no date).

REFERENCES

Bacharach, Samuel B., David B. Lipsky, and Joseph B. Shead, Paying for Better Teaching. Ithica: Organizational Analysis and Practice, Inc., 1984.

Calhoun, Frederick S. and Nancy J. Protheroe. Merit Pay Plans for Teachers: Status and Descriptions. Arlington, Va.: Educational Research Service, 1983.

The Catholic High School: A National Portrait. Washington, D.C.: National Catholic Education Association, 1985.

Coltham, Jeanette B., "Educational Accountability: A English Experiment and Its Outcome," School Review, (Nov. 1972), pp. 15-34.

Evendon, E.S. Teachers' Salaries and Salary Schedules in the United States, 1918-19. Washington, D.C.: National Education Association, 1918.

Gramlich, Edward and Patricia Koshel. Educational Performance Contracting. Washington, D.C.: Brookings Institution, 1975.

Hanushek, Eric, "Throwing Money at Schools," Journal of Policy Analysis and Management, 1(Fall 1981)1, pp. 19-41.

Hirschman, Albert O. Exit, Voice, and Loyalty. Cambridge, Ma.: Harvard University Press, 1970.

Johnson, Susan M., "Merit Pay for Teachers: A Poor Prescription for Reform," Harvard Educational Review, 54(May 1984)2, pp. 175-185.

The Nation Responds: Recent Efforts to Improve Education. Washington, D.C.: U.S. Department of Education, 1984.

National Education Association (NEA), Practices Affecting Teacher Personnel, Research Bulletin, Vol. 6, No. 4 (Washington, D.C., 1928.

National Education Association (NEA), Report of the Salary Committee, Teachers' Salaries and Salary Trends in 1923. Washington, D.C.: NEA, 1923.

Pencavel, John H., "Work Effort, On-the-Job Screening, and Alternative Methods of Remuneration," in (R. Ehrenberg, ed.), Research in Labor Economics, Vol.I, Greenwich, Ct.: JAI Press, 1977, pp.225-258.

Porwoll, Paul J. Merit Pay for Teachers. Arlington, Va.: Educational Research Service, 1979.

Seiler, Eric, "Piece Rate vs. Time Rate: The Effect of Incentives on Earnings," The Review of Economics and Statistics, 66(Aug. 1984)3, pp. 363-375.

Small Grants for Teachers. Pittsburgh: Allegheny Conference on Community Development, no date.

Williamson, Oliver E., Markets and Hierarchies: Analysis and Antitrust Implications. New York: Free Press, 1975.

Wise, Arthur E., Linda Darling-Hammond, Milbrey W. McLaughlin, and Harriet T. Bernstein. Teacher Evaluation: A Study of Effective Practices. Santa Monica, Ca., 1984.